## Карточка 6 «Алфавитно-частотный словарь»

1. Постройте программу, которая составляет алфавитно-частотный словарь для заданного файла *listwords.txt* со списком слов. Программа должна выводить результат работы в другой файл (*listwords.dic*).
2. Выведите все найденные слова в файл (*listwords2.dic*) в порядке убывания частоты, то есть в начале списка должны стоять слова, которые встречаются в файле *listwords.txt* чаще всех. Слова, встречающиеся одинаковое количество раз, должны быть расставлены по алфавиту.
3. Постройте программу, которая составляет алфавитно-частотный словарь для заданного файла (*text.txt*) с текстом. Не забудьте, что строки текста нужно разделить на слова и удалить все знаки препинания.

**Для тех, кто сдает ЕГЭ**

**Задача 1\*.** Текстовый файл 24-157.txt содержит только заглавные буквы латинского алфавита (ABC…Z). Определите символ, который реже всего встречается в файле. Если таких символов несколько, выведите тот, который стоит в алфавите последней.

**Задача 2\*.** Текстовый файл 24-s2.txt содержит только заглавные буквы латинского алфавита (ABC…Z). Определите символ, который чаще всего встречается в файле сразу после буквы A. В ответе запишите сначала этот символ, а потом сразу (без разделителя) сколько раз он встретился после буквы А. Если таких символов несколько, нужно вывести тот, который стоит раньше в алфавите. Например, в тексте ABCAABADDD после буквы A два раза стоит B, по одному разу – A и D. Для этого текста ответом будет B2

**Задача 3\*.** Текстовый файл 24-s2.txt содержит только заглавные буквы латинского алфавита (ABC…Z). Определите символ, который чаще всего встречается в файле между буквами X и Z, так что X стоит слева от него, а Z – справа. В ответе запишите сначала этот символ, а потом сразу (без разделителя) сколько раз он встретился между буквами X и Z. Если таких символов несколько, нужно вывести тот, который стоит позже в алфавите. Например, в тексте XBZCXXZXBZXDZDD между буквами X и Z два раза стоит B, по одному разу – X и D. Для этого текста ответом будет B2.

**Задача 4\***. Текстовый файл 24-s1.txt состоит не более чем из 106 заглавных латинских букв (A..Z). Текст разбит на строки различной длины. Необходимо найти строку, содержащую наименьшее количество букв A (если таких строк несколько, надо взять ту, которая в файле встретилась раньше). Определите, какая буква встречается в этой строке чаще всего. Если таких букв несколько, надо взять ту, которая стоит последней в алфавите. Запишите в ответе эту букву, а затем – сколько раз она встречается во всем файле.

Пример.Исходный файл:

**ZZQAQ**

**ZAVLAB**

**KRAKTU**

В этом примере в первой и третьей строках по одной букве A, во второй – две. Берём первую строку, т.к. она находится в файле раньше. В этой строке чаще других встречаются буквы Z и Q (по два раза), выбираем букву Z, т. к. она позже стоит в алфавите. В ответе для этого примера надо записать Z3, так как во всех строках файла буква Z встречается 3 раза.

**Задача 5\*.** Текстовый файл [24-164.txt](https://kpolyakov.spb.ru/cms/files/ege-sym/24-164.txt) состоит не более чем из 106 символов и содержит только заглавные буквы латинского алфавита (ABC…Z). Текст разбит на строки различной длины. В строках, содержащих менее 20 букв E, нужно определить и вывести максимальное расстояние между одинаковыми буквами в одной строке.

Пример. Исходный файл:

VOVA  
ZAEALE  
QRAEQT

В этом примере во всех строках меньше 20 букв E. Самое большое расстояние между одинаковыми буквами – в третьей строке между буквами Q, расположенными в строке на 1-й и 5-й позициях. В ответе для данного примера нужно вывести число 4.

**Задача 6\*\*.** Текстовый файл 2-6.txt состоит не более чем из 106 символов и содержит только заглавные буквы латинского алфавита (ABC…Z). Текст разбит на строки различной длины. Необходимо найти строку, содержащую самую длинную цепочку стоящих подряд одинаковых букв. Если таких строк несколько, надо взять ту, которая в файле встретилась раньше. Определите, какая буква встречается в этой строке чаще всего. Если таких букв несколько, надо взять ту, которая стоит раньше в алфавите. Запишите в ответе эту букву, а затем – сколько раз она встречается во всем файле.

Пример. Исходный файл:

ZZQABA

ZALAAC

QRAQUT

В этом примере в первой и второй строках наибольшая длина цепочек одинаковых буквы равна 2 (ZZ в первой строке, AA во второй), в третьей – 1. Берём первую строку, т.к. она находится в файле раньше. В этой строке чаще других встречаются буквы Z и A (по 2 раза), выбираем букву A, т. к. она стоит раньше в алфавите. В ответе для этого примера надо записать A6, так как во всех строках файла буква A встречается 6 раз.